View-based Shape Similarity using Mutual Information Spheres

Francisco González García, Miquel Feixas and Mateu Sbert

Universitat de Girona, Spain

Abstract

We present a new method for computing the shape similarity between 3D polygonal models using an information-theoretic viewpoint selection framework. Given a 3D model, a sphere of viewpoints surrounding this model is used to obtain its shape signature from the mutual information of each viewpoint. This signature represents the essence of the shape from a view-based approach. Then, in order to quantify the dissimilarity between two models, their mutual information spheres are registered by minimizing the L2 distance between them. Several experiments show the discrimination capabilities of our approach and its potential suitability for object recognition.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism

1. Introduction

Quantifying the *shape similarity* between 3D polygonal models is a key problem in different fields, such as computer graphics and computer vision. Many research has been done in these areas, however we only mention some of the most recent results. Funkhouser et al. [FMK*02] present methods for automatic shape-based retrieval of 3D models and a web-based search engine. Shilane et al. [SMKF04] describe a publicly database of polygonal models and a suite of tools for comparing shape matching and classification algorithms. Osada et al. [OFCD02] and Gal et al. [GSCO07] propose different methods for computing 3D shape signatures and their application to object classification and database retrieval. See [JT04] for a survey on content-based 3D shape retrieval.

In computer graphics, several *viewpoint quality* measures have been applied in areas such as image-based modeling [VFSH03] and volume rendering [BS05, VFSG06]. In object recognition, best view selection is also a fundamental task. Many works have demonstrated that the recognition process is view-dependent [BET95,TBZB97]. In [TBZB97], the authors found that "visual recognition may be explained by a view-based theory in which viewpoint-specific representations encode both quantitative and qualitative features".

In this paper, the shape similarity problem is tackled from an information-theoretic framework introduced in [VFSG06, FSG07]. Given a set of viewpoints surrounding a 3D model, we calculate its shape signature given by the viewpoint mutual information sphere. Then, mutual information spheres are registered by finding the minimum dissimilarity between them.

2. Basics: Viewpoint Mutual Information

In this section we review the definition of mutual information (MI) [CT91] applied to a viewpoint information channel. In [FSG07], a viewpoint selection framework has been constructed from an information channel $V \rightarrow O$ between the random variables V (input) and O (output), which represent, respectively, a set of viewpoints and the set of polygons of an object. This *viewpoint channel* is defined by a conditional probability matrix obtained from the projected areas of polygons at each viewpoint. Viewpoints are indexed by v and polygons by o, and capital letters V and O, used as arguments of $p()$, denote probability distributions. For instance, while $p(v)$ denotes the probability of a single viewpoint v, $p(V)$ represents the input distribution of the set of viewpoints.

The viewpoint channel can be interpreted as an *observa-*

Figure 1: *Models used in our experiments. From left to right: fishes, chairs, cars and shoes.*

tion channel where the conditional probabilities represent the probability of *seeing* a determined polygon from a given viewpoint. The three basic elements of this channel are:

- Conditional probability matrix $p(O|V)$, where each element $p(o|v)$ is defined by the normalized projected area of polygon o over the sphere of directions centered at viewpoint v. Conditional probabilities fulfil $\sum_{o\in\mathcal{O}} p(o|v) = 1$.
- Input distribution $p(V)$, which represents the probability of selecting a viewpoint and is obtained from the normalization of the projected area of the object at each viewpoint.
- Output distribution $p(O)$, defined by $p(o) = \sum_{v\in\mathcal{V}} p(v)p(o|v)$, which represents the average projected area of polygon o.

The *mutual information* between V and O, that expresses the degree of *dependence* or *correlation* between the set of viewpoints and the object, is given by

$$I(V,O) = \sum_{v\in\mathcal{V}} p(v) \sum_{o\in\mathcal{O}} p(o|v) \log \frac{p(o|v)}{p(o)} = \sum_{v\in\mathcal{V}} p(v) I(v,O),$$

where

$$I(v,O) = \sum_{o\in\mathcal{O}} p(o|v) \log \frac{p(o|v)}{p(o)} \qquad (1)$$

has been defined as the *viewpoint mutual information* (VMI), which represents the degree of dependence between the viewpoint v and the set of polygons, and it is a measure of the *quality* of viewpoint v. In this framework, the best viewpoint is defined as the one that has *minimum* VMI. One of the main properties of VMI is its robustness to deal with different discretisations of the model (see [VFSG06, FSG07]).

To compute VMI, we need to estimate the projected area of the visible polygons of the object at each viewpoint. Before projection, a different color is assigned to each polygon. The number of pixels with a given color divided by the total number of pixels projected by the object gives us the relative area of the polygon represented by this color (conditional probability $p(o|v)$). In our experiments, all the objects are centered in a sphere of 642 viewpoints built from the recursive discretisation of an icosahedron and the camera is looking at the center of this sphere. The VMI sphere is represented by a color map, where red and blue colors correspond respectively to the best and worst views.

3. View-based Shape Similarity

As we have seen in the previous section, VMI measures the degree of correlation between a viewpoint and the model.

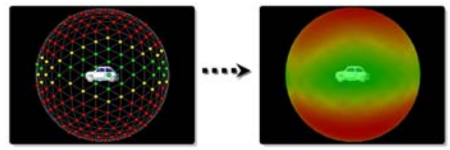

Figure 2: *The VMI sphere of the first car model shown in Figure 1. The VMI sphere on the right-hand side has been obtained by linear interpolation of the VMI values at viewpoint positions (left image).*

The VMI sphere (Figure 2) is obtained from the mutual information at each viewpoint and is now interpreted as a *shape signature* that captures the essence of the shape from a view-based approach. Then, the VMI spheres can be registered to determine the similarity between two models.

The goal of the registration between two VMI spheres is to find the transformation that brings one sphere (floating) into the best possible spatial correspondence with the other one (fixed) by minimizing a dissimilarity metric. The components of our registration method and their interconnections are shown in Figure 3. The basic input data to the registration process are two VMI spheres. The transform component represents the spatial mapping of points from the fixed sphere space to points in the floating sphere space. The interpolator component is used to evaluate floating sphere values at non-viewpoint positions and, finally, the metric module provides a measure of how well the fixed sphere is matched by the transformed floating sphere.

The steps followed by our method to achieve the best matching between the fixed and the floating sphere are:

1. **Interpolation**. The discrete nature of our VMI spheres implies the need of having an interpolator component. The nearest neighbor interpolator has been used. This means that when we need to evaluate values at non-viewpoint positions on the floating sphere we will use the VMI value of the closest viewpoint.
2. **Comparison**. To quantify the quality of the alignment between the fixed and the floating sphere we need a dissimilarity metric. In our method we have adopted the L2 distance between the VMI values of the spheres S_1 and S_2 corresponding to models O_1 and O_2, respectively:
$$D(S_1,S_2) = \sqrt{\sum_{v\in\mathcal{V}} (I(v,O_1) - I(v,O_2))^2}. \qquad (2)$$
3. **Transformation**. We need two transformation parame-

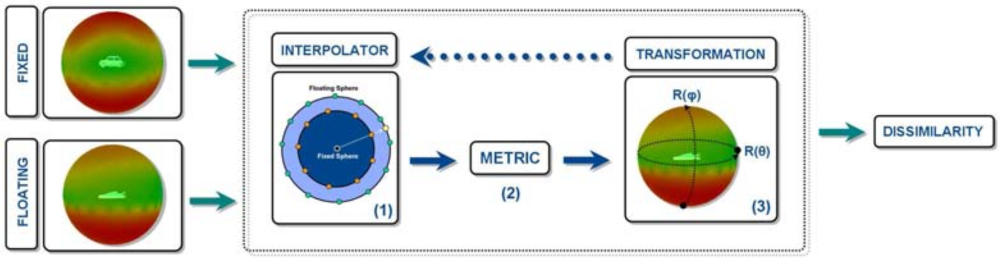

Figure 3: *Registration of VMI spheres and its main components.*

ters (degrees of freedom): R(θ) and R(φ), defined respectively as the rotation around Z and Y axis. These two parameters take values in the range [0°,360°] and [0°,180°], respectively.

When all the possible registration positions (dependent on the transformation parameters) have been analyzed, the correct matching is given by the minimum dissimilarity. In our current implementation, running on a Pentium IV 3GHz machine with 2GB RAM and an NVidia GeForce 8800 GTX, a single registration takes approximately two minutes when the transformation parameters are increased in steps of five degrees. The cost of this registration process could be considerably improved by using numerical optimizers. The memory space consumption required can be considered negligible.

4. Results

The view-based shape matching described in the previous section has been incorporated into our viewpoint software using the Ogre3D rendering engine (`http://www.ogre3d.org`). In our experiments, the viewpoint sphere is built from the smallest bounding sphere of the model. The radius of the viewpoint sphere is three times the radius of the bounding one.

In order to demonstrate the performance of our approach we have used four families of models (fishes, chairs, cars and shoes) where each one is composed by four different samples (Figure 1). Our registration method has been applied to all pairs of models obtaining the dissimilarity (2) between the VMI spheres. The transformation parameters R(θ) and R(φ) take values in intervals of five degrees.

From the dissimilarity values obtained with the spherical registration, we have built the dissimilarity map shown in Figure 4. Each row and column of the map represents an object and the color given to the intersection between them is the resulting dissimilarity (not intersected regions between models have been linearly interpolated). Red and blue colors represent dissimilar and similar objects, respectively. Note the blue regions along the diagonal as well as the predominance of warm colors while moving away. Let us note the large blue and green area surrounding the car family region and the dissimilarity of fishes with respect to the rest of the models.

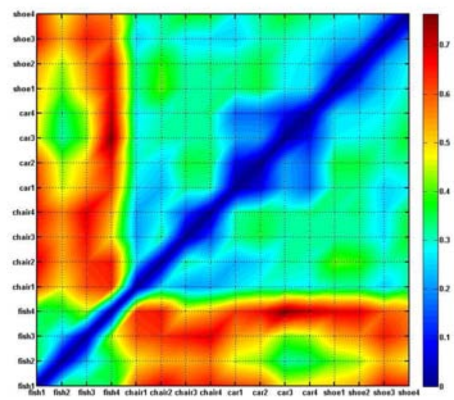

Figure 4: *Dissimilarity map. Blue and red values correspond to the most similar and dissimilar models respectively.*

In Figure 5 we show the shape similarity between the first model of each family and the rest of models. The list of models has been ordered according to the dissimilarity obtained with the spherical registration. Observe the perfect matching in the fish and car families. We also want to stand out the good behavior of the chairs and shoes.

5. Conclusions and Future Work

This paper is a first step in exploring the possibilities of an information-theoretic viewpoint selection framework to quantify the shape dissimilarity between 3D polygonal models. The presented approach is based on two phases. First, the viewpoint mutual information sphere is calculated for each model. This sphere is considered as a shape descriptor. Second, the registration between the mutual information spheres is performed by minimizing the dissimilarity between them. Several experiments show the potential suitability of our ap-

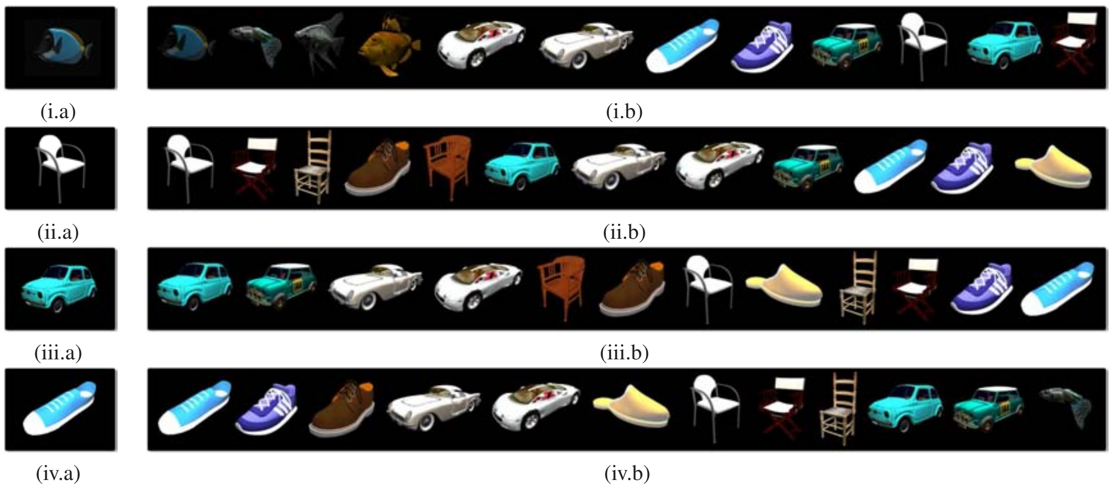

Figure 5: *In column (a) we show the first model of each family and in column (b) the list of the first twelve objects sorted by its similarity with respect to the target model (a).*

proach to pre-classify 3D polygonal models for object recognition. The initial results encourage us to explore the use of (1) stability and saliency spheres (see [FSG07]) as shape signatures in combination with VMI-spheres, (2) other metrics for spherical registration and dissimilarity quantification, (3) numerical optimizers to speed up the registration process, (4) several interpolators and (5) different resolutions of both the viewpoint sphere and the model mesh.

Acknowledgments

This project has been funded in part with grant numbers TIN2004-07451-C03-01 of the Spanish Government and IST-2-004363 (GameTools: Advanced Tools for Developing Highly Realistic Computer Games) from the VIth European Framework.

References

[BET95] BÜLTHOFF H., EDELMAN S., TARR M.: How are three-dimensional objects represented in the brain? *Cerebral Cortex 5* (1995), 247–260.

[BS05] BORDOLOI U. D., SHEN H.-W.: Viewpoint evaluation for volume rendering. In *Visualization, IEEE 2005* (May 2005), pp. 487–494.

[CT91] COVER T. M., THOMAS J. A.: *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[FMK*02] FUNKHOUSER T., MIN P., KAZHDAN M., CHEN J., HALDERMAN A., DOBKIN D., JACOBS D.: A search engine for 3d models. *ACM Transacions on Graphics* (2002).

[FSG07] FEIXAS M., SBERT M., GONZÁLEZ F.: *A Unified Information-Theoretic Framework for Viewpoint Selection and Mesh Saliency.* Research Report IIiA 07-03-RR, IIiA - Institut d'Informàtica i Aplicacions, Universitat de Girona (Girona, Spain), 2007. Submitted.

[GSCO07] GAL R., SHAMIR A., COHEN-OR D.: Pose-oblivious shape signature. *IEEE Transactions on Visualization and Computer Graphics 13*, 2 (2007), 261–271.

[JT04] J. TANGELDER R. V.: A survey of content based 3d shape retrieval methods. *Proceedings of Shape Modeling International* (2004), 145–156.

[OFCD02] OSADA R., FUNKHOUSER T., CHAZELLE B., DOBKIN D.: Shape distributions. *ACM Transaction on Graphics 21*, 4 (October 2002), 807–832.

[SMKF04] SHILANE P., MIN P., KAZHDAN M., FUNKHOUSER T.: The princeton shape benchmark. In *Shape Modeling International* (June 2004), pp. 167–178. Held in Genova, Italy.

[TBZB97] TARR M., BÜLTHOFF H., ZABINSKI M., BLANZ V.: To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science 8*, 4 (1997), 282–289.

[VFSG06] VIOLA I., FEIXAS M., SBERT M., GRÖLLER M. E.: Importance-driven focus of attention. *IEEE Trans. Vis. Comput. Graph. 12*, 5 (2006), 933–940.

[VFSH03] VÁZQUEZ P. P., FEIXAS M., SBERT M., HEIDRICH W.: Automatic view selection using viewpoint entropy and its application to image-based modeling. *Computer Graphics Forum* (Desember 2003).